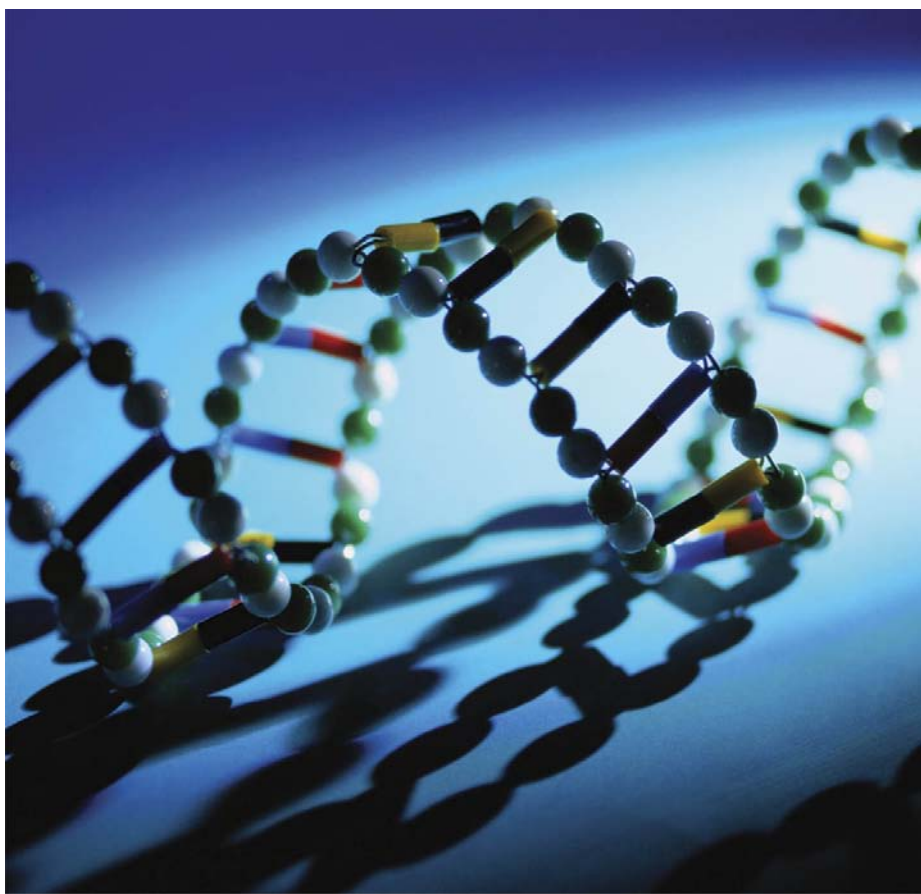
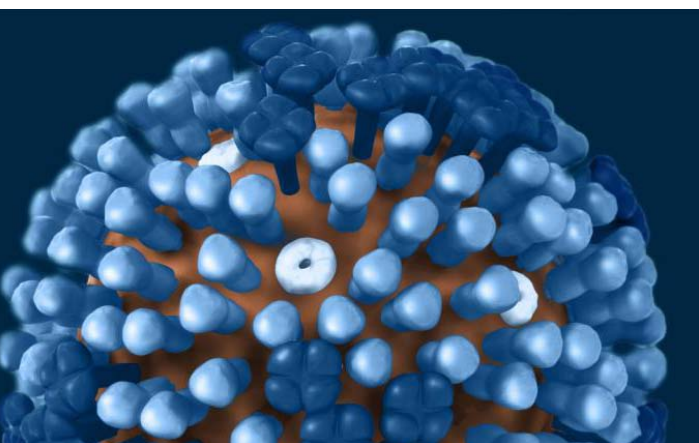


# Building on a Decade of Accomplishments

*Report of the 2010 Blue Ribbon Panel on Genomics*



Division of Microbiology and Infectious Diseases, Genomics Programs  
National Institute of Allergy and Infectious Diseases  
National Institutes of Health



# **Building on a Decade of Accomplishments: Report of the 2010 Blue Ribbon Panel on Genomics**

**Division of Microbiology and Infectious Diseases, Genomics Programs**  
National Institute of Allergy and Infectious Diseases  
National Institutes of Health

## **TABLE OF CONTENTS**

**Introduction**

**Background**

**Current NIAID/DMID Genomics Programs**

**The Blue Ribbon Panel**

**Guiding Principles and Recommendations**

## Introduction

In 2010, the Division of Microbiology and Infectious Diseases (DMID), National Institute of Allergy and Infectious Diseases (NIAID) convened a ‘Blue Ribbon Panel’ to review its Genomics Programs and provide recommendations for future activities. Throughout the Blue Ribbon Panel meeting, several important themes emerged. Based on these themes, the Panel provided overarching guiding principles as well as more specific recommendations as the Programs move into the next decade. Panelists were specifically instructed to cast a very wide net in their deliberations, using the broadest definition of genomics.<sup>1</sup> The focus of the meeting also included review of computational infrastructure, future technologies, and policies for data release and sharing.

## Background

Over the last 10 years, genomics has powered a quantum leap forward in the study of infectious diseases. Genomics technologies have emerged as potent approaches for understanding how microbes cause disease, as well as the interaction between pathogen and host. Thousands of microbial genomes have been sequenced, providing a wealth of data on pathogenic and non-pathogenic organisms. This information, combined with sequence information on the human genome and the human microbiome, provides a valuable resource to help scientists identify the functions of specific microbial DNA sequences and shed light on microbial biology and pathogenesis.

Genomics research has already generated new strategies for diagnosing, preventing, and treating infectious diseases. Analyzing human genome sequences is enhancing our understanding of host response and variation in individual susceptibility to microbial pathogens and infectious diseases. These efforts may also provide insights about individual responses to vaccines and therapeutics.

In recent years, the field of metagenomics has emerged as a complementary approach to studying microbes. By studying microbial communities inhabiting a particular human body site, metagenomics is enabling scientists to gain insight into their role in health and disease without culturing individual microbes.

NIAID’s Division of Microbiology and Infectious Diseases is a pioneer in using genomics technologies to study infectious diseases, supporting successful sequencing of thousands of microbial genomes including *influenza*, *dengue*, *Bacillus anthracis*, *Plasmodium falciparum*, *Aspergillus fumigatus*, *Mycobacteria tuberculosis*, and *Staphylococcus aureus*. Today, sequencing a bacterial genome costs less than a dollar, yet analysis may cost as much as tens of thousands of dollars.<sup>2</sup> DMID has been expanding genomics activities over the last decade to provide the scientific community with genomic data as well as resources such as reagents, databases, software, and computational tools that are essential for analyzing and applying research findings. Data and resources generated through the genomics initiatives are rapidly made available to the scientific community.

---

<sup>1</sup> Throughout this document, the term genomics is used broadly to include functional genomics, proteomics, bioinformatics, computational biology, structural genomics, lipidomics, metabolomics, and glycomics. Where applicable, systems biology is included as well.

<sup>2</sup> Noted at Blue Ribbon Panel meeting by Panel member, Dr. Eric Lander.

Sequence data, combined with other biochemical and microbiological information, is being used to improve detection of pathogens, diagnose infectious diseases, and identify potential new targets for novel drugs and vaccines. In addition, comparing the sequences of different strains, species, and clinical isolates is crucial for identifying genetic polymorphisms that correlate with phenotypes such as drug resistance, morbidity, and infectivity.

## Current NIAID/DMID Genomics Programs

The current Genomics Programs (Figure 1) provide comprehensive genomics, functional genomics, proteomics, structural genomics, bioinformatics and other ‘omics’ resources and reagents to the scientific community for basic and applied research in infectious diseases. The ultimate goal of the Programs is to facilitate understanding of biology of the pathogen, pathogenesis, pathogen-host interaction, and to develop potential new targets and platforms for discovering novel drugs, vaccines and diagnostics. The components of the Genomics Programs share certain characteristics: they probe the enormous complexity and diversity of biological systems; they use techniques that can examine many thousands of separate biological entities or phenomena in a single experiment; and they generate new data on an unprecedented scale, all of which must be stored, curated, analyzed, and shared to the greatest extent possible.

The Genomics Programs include:

- **Genome Sequencing Centers** provide human and microbial genotyping as well as rapid and cost-efficient production of high-quality genome sequences of human pathogens and invertebrate vectors of infectious diseases.
- **Pathogen Functional Genomics Resource Center** provides genomic resources, data sets, and reagents to elucidate the functions of specific genes and proteins.
- **Proteomics Research Centers** characterize the proteomes of pathogens and/or host cells; identify proteins associated with the biology of microbes, mechanisms of microbial pathogenesis, and host response to infection.
- **Clinical Proteomics Centers for Infectious Diseases and Biodefense** discover, qualify, and verify candidate protein biomarkers with potential to be used clinically to improve diagnosis, determine susceptibility to infection, and monitor responses to drugs and vaccines.
- **Structural Genomics Centers for Infectious Diseases** experimentally characterize the three-dimensional atomic structure of proteins of pathogenic and non-pathogenic microbes.
- **Systems Biology Centers for Infectious Disease Research** identify and analyze molecular interaction networks and predictive models of microbial pathogens and their host cells through a combination of computational and experimental high-throughput technologies.
- **Bioinformatics Resource Centers** collect, integrate, and provide open access to research data on microbial organisms and vectors of infectious diseases in a user-friendly format. Develop and share open-source software tools; provide bioinformatics services and training to the scientific community.

For more information about these Programs, visit [www.niaid.nih.gov/topics/pathogenGenomics](http://www.niaid.nih.gov/topics/pathogenGenomics).

# NIAID/DMID Genomics Programs - 2011

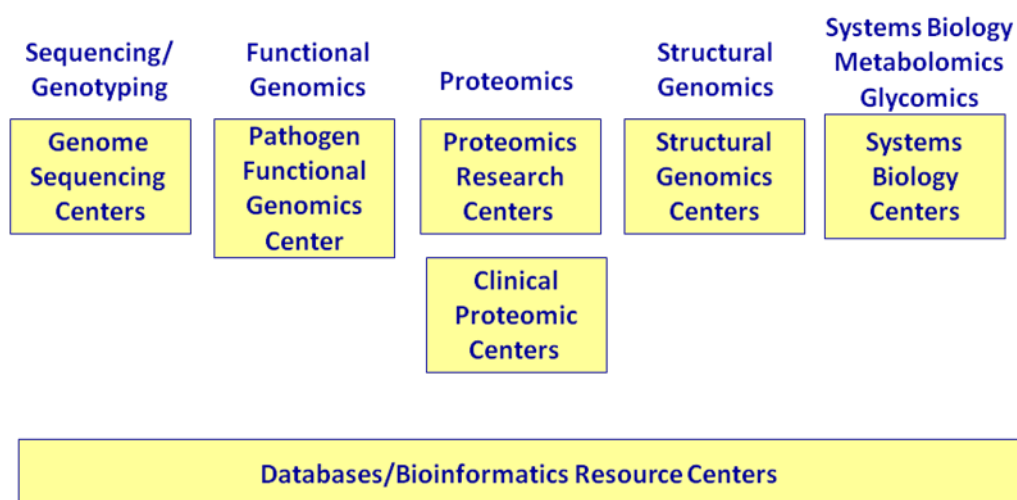


Figure 1. The Genomics Programs of NIAID's Division of Microbiology and Infectious Diseases support a comprehensive research portfolio that is helping to increase understanding of infectious disease pathogens and host-pathogen interaction, including the human microbiome. Fostering innovative technologies and providing essential tools for analyzing and managing genomic information are important components of the Programs. Ten years after the Programs' inception, the concept of genomics has broadened to include many other 'omics' fields, setting the stage for discussions about future directions.

## The Blue Ribbon Panel

The field of genomics has advanced rapidly in the decade since DMID initiated its Genomics Programs. Thus, a multidisciplinary team of experts was convened to review the current Programs, and provide recommendations for the future. The Blue Ribbon Panel consisted of more than 50 experts in infectious diseases, genomics, bioinformatics, proteomics, other 'omics,' computational biology, and software engineering. Participants included representatives from industry, academia, non-profit organizations, and experts from across NIH as well as other Federal agencies. Dr. Claire Fraser-Liggett of the University of Maryland School of Medicine and Dr. Michael Osterholm of the University of Minnesota served as Panel co-chairs. The panel assembled for a one-day meeting on February 4, 2010. (See Appendix I for meeting agenda and Appendix II for a list of panelists.)

Presentations from invited speakers encouraged discussion and deliberation. For the review of current Programs, panelists worked in breakout groups with representatives from different disciplines. They focused on four areas related to the provision of resources to the scientific community: diversity; value; accessibility; and integration of services, reagents, and resources. Each breakout group later presented key points to the larger group. The same breakout groups convened to discuss possible future activities for the Programs and provide recommendations for consideration as the Genomics Programs move into the next decade of using genomic technologies to study infectious diseases. A final session summarized specific recommendations.

## Guiding Principles and Recommendations

In reviewing the last decade of the Genomic Programs, panelists recognized the important efforts and achievements of the Programs and indicated that the enterprise has substantially accelerated scientific progress in infectious diseases. They commended DMID on the range of services provided and the ability to balance competing priorities to address a variety of diverse research needs and communities. Panelists also noted the Programs' leadership role in setting guidelines for the domestic and international infectious diseases and genomics communities to promote rapid yet judicious release of genomics data sets and reagents. Panelists recognized the range of services offered by the Genomics Programs, including:

- A broad spectrum of genomics data sets, reagents and resources.
- Computational and data analysis tools and bioinformatics services.
- A flexible infrastructure to respond to needs of community and public health emergencies.
- New, state-of-the-art technologies.
- Enhanced quality control of reagents and resources generated.

The panel outlined four guiding principles and associated recommendations that should underpin the Genomics Programs as they move into the next decade.

### **1. Set ambitious and challenging goals for applying genomic approaches and technologies to increasing understanding of infectious diseases.**

Scientific opportunities in microbiology and immunology—powered by advances in genomic technologies—are simply extraordinary. Projects that would have been called impossible even a few years ago now seem possible. That fact, however, means that NIAID/DMID, working in collaboration with scientific partners, must be very ambitious in its goals and programs to support the application of genomic technologies to understanding infectious disease pathogens at a systems level. This includes pursuing large-scale, temporal and spatial studies of infectious diseases that characterize the pathogen, pathogen-host interaction, and host response in real time as diseases progress.

- *Reconstruct and characterize entire human microbial communities, especially those consisting of multiple species.*
- *Characterize variation among microorganisms that influence health.*
- *Pursue studies to understand uncharacterized genomic regions in microbes.*
- *Profile the immune system in health and illness.*

### **2. Promote the use of cutting-edge genomic technologies and integrated data sets to establish new strategies for developing drugs, vaccines and diagnostics for infectious diseases.**

The last decade has seen breakthroughs in genomic technologies that are powering ground-breaking advances in microbiology, immunology, and infectious diseases. The Programs should continue to assist scientists to manage, analyze, and share high-quality data and should continue to make advanced genomic technologies accessible to the research community. Applying knowledge gained to clinical practice should be a high priority, including providing new molecular targets, signatures, and platforms to develop novel drugs, vaccines and diagnostics for infectious diseases. Setting ambitious goals and undertaking creative and innovative strategies could have an enormous impact on preventing and treating infectious diseases globally.

- *Continue to use systems biology strategies to understand infectious diseases and identify and reconstruct entire networks of metabolic pathways of pathogens, host responses, and the complex interplay between them.*
- *Continue to generate new targets using innovative structural genomics and sequencing approaches for drug discovery.*
- *Use genomic technologies to develop platforms for improved diagnostics and vaccines that can be used in clinical settings.*
- *Validate biomarkers and molecular signatures that have clinical applications.*
- *Integrate potential for genomic studies into clinical trials.*
- *Continue to generate data sets that can be used to enhance public health surveillance and response.*

### **3. In partnership with the scientific community, continue to support bioinformatic and computational infrastructure that can adapt to evolving research needs.**

Data are being generated at an enormous rate need to be organized and analyzed. The Programs must therefore continue to provide the bioinformatics infrastructure, computational tools, and organizational methods that scientists need now, and anticipate what will be needed in five to ten years as even larger and more complex data sets are generated. These efforts must not be stand-alone endeavors, but rather integral parts of the Centers' work. This will facilitate partnerships within the scientific community and sharing of expertise, infrastructure, and costs among industry and the institutions that perform the research and generate the data. DMID should take advantage of new approaches to data storage and analysis including cloud computing and other next-generation computational strategies.

- *Using the current Bioinformatics Resource Centers as a model, build 'with and for' the scientific community in order to clearly understand and meet the need for bioinformatics and computational tools.*
- *Support data management, visualization, integration, and analysis as well as other needs to analyze diverse data sets.*
- *Support development of computational tools and training needed for data management and analysis.*  
*Develop tools to use large data sets to establish predictive models of molecular pathways of pathogen host interaction.*

### **4. Continue to play a leadership role in setting guiding principles for rapid data and reagent release.**

Because data generated by the Genomics Programs constitutes a very large percentage of available genomic data, it is essential that it be made rapidly available to the scientific community. Establishing guiding principles for rapid, yet judicious, data release will advance the ultimate goal of the Genomics Programs: to further the pursuit of new discoveries about the causes, treatment, and ultimate prevention of infectious and immune-mediated diseases.

- *Cultivate partnerships with the scientific community to promote rapid yet judicious release of data.*
- *Ensure that publicly released data sets from the Genomics Programs are of highest quality and utility.*
- *Generate clear, concise data release guidelines and policies.*
- *Implement guidelines and policies consistently and promote fair use of data.*